

Participación Educativa

REVISTA DEL CONSEJO
ESCOLAR DEL ESTADO

Ministerio
de Educación, Cultura
y Deporte

Transparencia y mejora de la educación

Consejo
Escolar
del Estado

Segunda Época/Vol. **2**/N.º **3**/2013



PARTICIPACIÓN EDUCATIVA

SEGUNDA ÉPOCA. NÚMERO 3. DICIEMBRE 2013

TRANSPARENCIA Y MEJORA DE LA EDUCACIÓN

ÓRGANOS DE DIRECCIÓN

Consejo de Dirección

Presidencia

Francisco López Rupérez
Presidente del Consejo Escolar del Estado

Vicepresidencia

María Dolores Molina de Juan
Vicepresidenta del Consejo Escolar del Estado

Secretario

José Luis de la Monja Fajardo
Secretario General del Consejo Escolar del Estado

Vocales

Montserrat Milán Hernández
Consejera de la Comisión Permanente

Roberto Mur Montero
Consejero de la Comisión Permanente

Jesús Pueyo Val
Consejero de la Comisión Permanente

Consejo de Edición

María Dolores Molina de Juan
(Consejo Escolar del Estado)

José Luis de la Monja Fajardo
(Consejo Escolar del Estado)

Isabel García García
(Consejo Escolar del Estado)

Almudena Collado Martín
(Consejo Escolar del Estado)

Antonio Frías del Val
(Consejo Escolar del Estado)

Juan Ramón Villar Fuentes
(Consejo Escolar del Estado)

Juan Luis Cordero Ceballos
(Consejo Escolar del Estado)

Consejo Asesor

Bonifacio Alcañiz García

Francisco J. Carrascal García

Julio Delgado Agudo

José Antonio Fernández Bravo

Mariano Fernández Enguita

Alfredo Fierro Bardají

José Luis Gaviria Soto

Samuel Gento Palacios

M^a Luisa Martín Martín

José María Merino

Sara Moreno Valcárcel

Arturo de la Orden Hoz

Francesc Pedró i García

Beatriz Pont

Gonzalo Poveda Ariza

M^a Dolores de Prada Vicente

Ismael Sanz Labrador

Rosario Vega García

Fotografías:

<http://bit.ly/1bwWS45>

ISSN 1886-5097

NIPO 030-13-166-8

ntic.educacion.es/cee/revista

participacioneduca@mecd.es



Presentación

Francisco López Rupérez **3**

Entrevista

José Manuel Romay Beccaría, *Presidente del Consejo de Estado* **5**

La transparencia: el marco político y jurídico

Transparencia y calidad de la democracia. Eugenio Nasarre Goicoechea **9**

La transparencia en la Unión Europea. Una visión comparada. M^a. Del Valle Ares González **15**

La transparencia en el marco normativo español. Oscar Sáenz de Santamaría Gómez-Mampaso **25**

La "Ley de transparencia, Acceso a la Información Pública y Buen Gobierno", y la responsabilidad en la acción pública. Pedro Ramón Gómez de la Serna **35**

La transparencia y la mejora de los sistemas educativos

La transparencia y la mejora de los sistemas educativos. El caso de PISA. Andreas Schleicher y Pablo Zoido Lobatón **45**

El Monitor de Educación y Formación en la Unión Europea. Joaquín Martínez Muñoz **53**

Los sistemas nacionales de indicadores como instrumentos de transparencia. Limitaciones y posibilidades. Felipe Martínez Rizo **67**

Evaluación de la eficacia de las políticas educativas y transparencia: la importancia de los experimentos aleatorizados. José García Montalvo **75**

Ámbitos de transparencia de los centros educativos

La web de centro como instrumento de transparencia educativa. Ferran Ruiz i Tarragó **83**

Un sistema básico de indicadores para los centros escolares. Julio Puente Azcutia **93**

El Expediente Educativo Electrónico. Hacia un sistema digital compartido para la gestión de información académica del alumno. José Quirino Vargas Ibáñez **101**

La transparencia económica en los centros públicos. Contextos, costes y eficiencias. José Manuel Bolado Somolinos **117**

La transparencia económica en los centros privados concertados. Contextos, costes y eficiencias. Luis Centeno Caballero **131**

Buenas prácticas y experiencias educativas

Spacial Helpers, el emprendimiento desde la escuela. Alicia Larráyoiz Echarte y Andrea Etxebarria Barinagarrementeria **141**

Propuesta bilingüe para el currículo integrado en las actividades extraescolares de las Secciones Españolas en el exterior. Experiencia en Liverpool. Cristina España Pérez **151**

Otros temas

Antecedentes, selección, formación y calidad de los directores escolares. Isabel Cantón Mayo **165**

Localización de fuentes cerebrales en niños generadas por estimulación multilingüística simultánea. Tomás Ortiz Alonso, Caetana Varela Hall y José Matías Santos García **175**

Recensiones de libros

"La escuela que necesitamos" (E.D. Hirsch. 2012). Inmaculada Egido Gálvez **187**

"Neuroeducación" (Francisco Mora. 2013). Tomás Ortiz Alonso **191**

"Catálogo bibliográfico de los siglos XVI al XIX. Instituto Isabel la Católica (Madrid)" (Encarnación Martínez Alfaro y Ángeles Castilla López). Juan Leal Pérez-Chao **194**

"International Handbook of Psychology in Education" (Littleton, K., Wood, C. y Staarman, J. K. 2012). Gonzalo Vázquez Gómez **196**

EVALUACIÓN DE LA EFICACIA DE LAS INTERVENCIONES EDUCATIVAS Y TRANSPARENCIA: LA IMPORTANCIA DE LOS EXPERIMENTOS ALEATORIZADOS¹

EVALUATION OF THE EDUCATIONAL INTERVENTIONS EFFICACY AND TRANSPARENCY: THE IMPORTANCE OF RANDOMIZED EXPERIMENTS¹

José García Montalvo

Universitat Pompeu Fabra

Resumen

El diseño experimental es considerado en muchas disciplinas como el procedimiento más adecuado para la evaluación del impacto de proyectos y programas. Este es el caso de la aprobación de nuevos medicamentos o innovaciones en procedimientos médicos. La evaluación educativa solía utilizar también procedimientos experimentales hasta sumirse en la oscuridad científica hacia mitad de los años 70. En este artículo se argumenta que es posible, y conveniente, realizar evaluaciones basadas en diseños experimentales en el campo de la educación en lugar de argumentar usando justificaciones apriorísticas o ideológicas. La gran ventaja de la aproximación experimental a la evaluación de las intervenciones educativas es la transparencia y credibilidad de sus resultados. Países como Estados Unidos o el Reino Unido están avanzando en la evaluación científica de las intervenciones educativas hasta el punto de mantener bases de datos con los resultados de experimentos sobre diferentes tipos de proyectos y programas.

Palabras clave: diseño experimental, causalidad, tamaño de las clases, calidad del profesor, evaluación, transparencia.

Abstract

The randomized control trial experimental design is the gold standard for the evaluation of projects and programs in many fields. This is the case of the approval of new drugs or the introduction of innovations in medical procedures. Up to the middle of the 70's there were also many experiments for the evaluation of educational interventions. Unfortunately this approach to the evaluation of education lost steam during many years in favor of aprioristic approaches and ideological biases. In this article I argue that it is possible, and convenient, to go back to the experimental design in the evaluation of educational projects and programs. The randomized control trial experiments have two basic advantages: they deliver credible and transparent results. The new research on the evaluation of educational projects and interventions in the US and the UK is moving towards the experimental design to the point of keeping databases with the results of the evaluation on many educational interventions.

Keywords: randomized control trial experiment, causality, class size, teacher's quality, evaluation, transparency.

“El diablo que hay en el mundo proviene de la ignorancia, y las buenas intenciones pueden hacer tanto daño como las malas intenciones, si falta el conocimiento”

Albert Camus, La Plaga

1. Introducción

Durante décadas las sucesivas reformas educativas que se han realizado en muchos países se han basado en modas pedagógicas y apriorismos muy poco adecuados al tratamiento científico de una disciplina tan importante como la educación². La evaluación basada en la evidencia ha estado ausente. ANGRIST (2004) plantea una pregunta fundamental: ¿Por qué en medicina o epidemiología los experimentos aleatorizados son la práctica habitual para realizar evaluaciones de impacto, mientras que en el campo educativo durante los últimos 30 años han estado ausentes? Si se quiere obtener la aprobación para la comercialización de una medicina basada en una nueva molécula que sirve de componente activo es obligatorio en muchos países realizar un experimento aleatorizado³ que debe proporcionar información sobre tres as-

pectos. Primero, que dicha molécula tiene efecto sobre un problema médico (tamaño de un tumor, recuento vírico, etc.). Segundo, que el compuesto es más efectivo que los existentes en el momento que se solicita su aprobación. Y tercero, que no tiene efectos secundarios perjudiciales, o que los mismos son muy limitados respecto a los beneficios que proporciona. De esta forma, el regulador en materia sanitaria concede a los experimentos aleatorizados el estatus de procedimiento de referencia en la generación de resultados transparentes y científicos. Precisamente la gran ventaja de los experimentos aleatorizados es que permiten realizar una atribución causal creíble y transparente del impacto de un determinado programa. Además, los resultados son fáciles de comunicar a los gestores políticos y la ciudadanía en general, por contraposición a muchos sofisticados procedimientos estadísticos, que son necesarios cuando se dispone solamente de datos observacionales.

Sin embargo, en muchos países cuando se decide financiar nuevos programas educativos o realizar una profunda reforma educativa, no se realiza una evaluación de su impacto. Peor aún, en ocasiones se toman decisiones que son contrarias a la evidencia empírica existente. ANGRIST (2004) argumenta que los polí-

¹ Artículo preparado para la revista Participación Educativa, vol. 2, número 3, Diciembre 2013. El autor agradece la financiación del proyecto SEJ2007-64340 del Plan Nacional de I+D del Ministerio de Economía y Competitividad.

² El New York Times publicaba el 2 de septiembre un artículo titulado “Guesses and hype give way to data in study of education”.

³ El caso de la Food and Drug Administration de Estados Unidos es paradigmático: solo acepta como procedimiento de validación de la nueva molécula la utilización de un experimento aleatorizado. Entendemos por un

experimento aleatorizado aquel diseño que trabaja con un grupo de control y un grupo tratado con asignación aleatoria de los individuos a cada grupo.

ticos que deben decidir sobre la implantación de nuevos programas educativos consideran que su impacto se produce a muy largo plazo⁴ mientras el ciclo político suele durar cuatro años. Si un medicamento produce efectos secundarios mortales la opinión pública rápidamente pedirá responsabilidades. Obviamente, un sistema pedagógico ineficaz o un currículum mal diseñado no pueden causar la muerte a los alumnos. Aunque sin duda cercena



su potencial de desarrollo futuro, su contribución al bienestar social y, en definitiva, la capacidad de mejorar el nivel de vida de la colectividad.

Lo más sorprendente de esta “fobia científica”⁵ es que gran parte de la comunidad pedagógica solía utilizar experimentos y procedimientos cuantitativos como método de análisis durante los 60 y parte de los 70. El famoso experimento de preescolar Perry de 1962 dio lugar al programa Head Starts, una intervención educativa masiva que recientemente ha sido también evaluada. Sin embargo, hacia mitad de los años 70 la investigación educativa decidió sumergirse en una fase oscuridad científica⁶. Algunos justifican el evitar los experimentos randomizados en educación en la existencia de unos supuestos problemas éticos: la injusticia que supone que algunos alumnos reciban una educación mejor que otros. Pero, a priori, no es posible saber si un programa tendrá éxito o no, al igual que un nuevo medicamento. Por tanto, pertenecer al grupo de control no supone una desventaja evidente. La selección por un procedimiento puramente aleatorio también evita cualquier suspicacia sobre una intención de favorecer a algunos alumnos o perjudicar a otros. Además, existen nuevas técnicas estadísticas que permiten que los estudiantes puedan negarse a participar en el experimento y, aún así, obtener resultados científicamente robustos del impacto de un determinado programa. Finalmente, muchas políticas y proyectos educativos tienen una aplicación en fases, lo que permite evitar por completo cualquier oposición que tenga como fundamento el perjuicio que se causa a los alumnos del grupo de control.

⁴ Aunque el efecto de mejorar la educación puede ser enorme en términos del aumento del nivel de vida de los ciudadanos. Ver HANUSHEK y WOESSMANN (2008, 2010).

⁵ Según definición de Cook (2001).

⁶ Hay alguna notable excepción en esta situación sombría como el experimento sobre la efectividad de las clases pequeñas conocido como proyecto STAR (Student-Teacher Achievement Ratio) que fue realizado entre 1985 y 1989.

No obstante, la educación como disciplina científica está retornando en algunos países y, particularmente en Estados Unidos, donde el Instituto de Ciencias de la Educación, dependiente del Ministerio de Educación de dicho país ha realizado 175 experimentos randomizados desde 2002⁷. La aprobación de la ley “No Child Left Behind” de 2001 ha sido un importante revulsivo para la vuelta a la investigación experimental. El articulado de la ley

requiere específicamente que la política educativa tenga fundamentos en investigación con base científica, definida como el uso de diseños metodológicos rigurosos y técnicas, incluyendo la creación de grupos de control y la asignación aleatoria, si fuera posible. Un aspecto muy importante de estas evaluaciones es la puesta a disposición de la comunidad educativa de los resultados obtenidos por estos experimentos aleatorizados. El Instituto de Ciencias de la Educación de EE.UU. mantiene una base de datos llamada “What Works clearinghouse” donde se sintetiza la evidencia disponible sobre la eficacia de programas, libros de texto, etc. basada en los experimentos aleatorizados disponibles. En una línea parecida el Education Endowment Fund mantiene una base de datos con meta-análisis de estudios sobre la efectividad de tres docenas de intervenciones educativas para las que existen evaluaciones aleatorizadas⁸. Además de los experimentos sociales existen otras

metodologías cuantitativas para aproximarse el impacto de un programa o un sistema educativo, que no utilizan una aproximación experimental⁹. Por ejemplo, la información que generan grandes estudios cuantitativos como PISA, TIMMS, PIRLS, etc. sin ser experimental, proporciona información muy valiosa que permite discutir sobre los méritos de distintos sistemas educativos. Sin embargo, es difícil en ocasiones extraer conclusiones robustas sobre los factores que determinan el éxito en estas pruebas estandarizadas internacionales. Por ejemplo, los estudiantes de Finlandia y Corea tienen puntuaciones muy altas en varias de las pruebas internacionales. Sin embargo, sus sistemas educativos son muy diferentes en múltiples dimensiones. La ventaja del método experimental es que puede medir el impacto de programas e intervenciones educativas de forma robusta y creíble y, a su vez, puede descomponer el éxito de los alumnos de uno u otro sistema en el efecto de sus componentes básicos. En todo caso, siempre es preferible tener la información de estas pruebas internacionales que mantener en la oscuridad o, peor aún, en el campo de la ideología el debate sobre cómo mejorar un sistema educativo y cuáles son las características de los sistemas educativos más efectivos. De hecho, en la situación actual del debate de la educación en muchos países, cualquier procedimiento que genere conocimiento basado en evidencia empírica debe ser saludado como un gran avance.

⁷ En la Unión Europea esta tendencia ha comenzado hace relativamente poco tiempo (GOGUEN y GURGAND 2012) si bien es cierto que siempre ha habido grupos de investigación cuantitativa en educación en departamentos de muchas universidades.

⁸ <http://educationendowmentfoundation.org.uk/toolkit/>

⁹ En este artículo se utiliza el término experimental como sinónimo de experimentación randomizada y pesado-experimentación basada en diseños como por ejemplo variables instrumentales, experimentos naturales o regresiones discontinuas.

2. ¿Qué sabemos sobre lo que funciona para mejorar el conocimiento de los alumnos?

La extensión de este trabajo impide llevar a cabo una revisión exhaustiva de la evidencia existente sobre qué funciona para mejorar el aprendizaje de los estudiantes. Pero, merece la pena centrarse al menos en algunos temas, para mostrar la importancia de una evaluación bien realizada y la potencia del método experimental en las ciencias sociales y, en concreto, en el campo de la educación. Los temas seleccionados son el efecto del tamaño de la clase, la importancia de la calidad del profesor y el impacto de las intervenciones tempranas.

2.1. El impacto del tamaño de la clase

Utilizando datos agregados por países, numerosos trabajos han mostrado que los conocimientos adquiridos por los estudiantes están, básicamente, no correlacionados con el gasto por estudiante o el gasto sobre el PIB de cada uno de los sistemas educativos. Estos estudios son interesantes para contrarrestar la falacia simplista que da por seguro que más recursos producen mejores resultados educativos. Sin embargo, los estudios macro tienen una capacidad muy limitada para extraer conclusiones sobre la dirección de causalidad entre estos dos factores. Por suerte existen muchos estudios basados en datos micro y experimentos randomizados, que analizan diversos aspectos del impacto de la cantidad de recursos en los resultados de los estudiantes. Quizás el recurso más estudiado sea la influencia del tamaño de la clase. Este es un asunto que ha interesado a los investigadores en el campo de la educación durante muchos años. ROCKOFF (2009) muestra que antes de 1940 ya se habían publicado 45 estudios sobre el impacto del tamaño de la clase sobre los resultados académicos de los estudiantes. De estos más de la mitad (24) se realizaron con un diseño experimental¹⁰. La presión financiera sobre las escuelas, por el sustancial incremento de la matriculación en las primeras décadas del siglo XX, fue determinante en la gran cantidad de estudios sobre el impacto del tamaño de la clase en esa época.

Pero sin duda el experimento más importante sobre el efecto del tamaño de la clase en el aprendizaje de los estudiantes es el STAR ("Student/Teacher Achievement Ratio) realizado en Tennessee entre 1985 y 1989. El análisis estudia una cohorte de 11.571 niños organizados en 79 colegios. Se hizo un seguimiento desde el último año de guardería hasta el tercer curso de primaria. Los profesores y los estudiantes fueron asignados aleatoriamente a clases dentro de cada colegio. Se analizaron dos tratamientos: clases pequeñas (15 alumnos), frente a clases grandes (22 alumnos); y clase grande con un profesor de apoyo, frente a clase grande sin profesor de apoyo. Los estudiantes se mantuvieron con el mismo tratamiento durante todos los años del experimento. Los resultados del experimento muestran que hay un efecto causal del tamaño de la clase, la calidad de profesor y los conocimientos de los compañeros de clase sobre los resultados de las pruebas cognitivas. Por el contrario la utilización de profesores de apoyo no tiene ningún efecto¹¹. Los estudiantes de las clases pequeñas tuvieron un resultado 0,22 desviaciones estándar mejor que los estudiantes de clases grandes después de 4 años (KRUEGER 1999). No obstante, esa mejora se concentra en el primer año (último año de guardería) y se va reduciendo a medida que avanzan en primaria. Por tanto, el efecto sobre los resultados en pruebas cognitivas es pequeño y decreciente. Además, los efectos de

estar en una clase pequeña eran mayores en niños afroamericanos de familias con pocos recursos económicos. KRUEGER (1999) estima una tasa de rentabilidad interna de la intervención del 6%.

Recientemente CHETTY y otros (2012) presentan una evaluación del efecto de haber asistido a una clase pequeña del STAR, sobre la situación laboral de los participantes en el estudio una vez que ya se habían incorporado al mercado laboral. Lo más sorprendente de la evaluación de CHETTY y otros (2012) es haber conseguido el emparejamiento de los datos de los alumnos de las escuelas del STAR con los datos de la hacienda pública de Estados Unidos. Los autores consiguieron enlazar el 95% de los alumnos del STAR con sus correspondientes declaraciones de renta entre 2005 y 2007. Esto significa que los estudiantes tenían 27 años cuando se efectúan las mediciones de renta. Los resultados de CHETTY y otros (2012) muestran que los resultados de las pruebas en el último año de guardería están muy correlacionados con los salarios, la asistencia a la universidad o la propiedad de una vivienda. El tamaño de la clase no afecta significativamente al salario (el estimador tiene una desviación estándar grande) aunque aumenta la probabilidad de asistir a la universidad.

2.2. La importancia de la calidad del profesor

Parece una verdad de Perogrullo, pero la calidad del profesor es el elemento más importante para el aprendizaje de los alumnos. Este factor es prácticamente el único que siempre resulta significativo estadísticamente en los estudios científicos sobre los factores determinantes del aprendizaje de los estudiantes. Aunque no hay acuerdo unánime sobre cómo se debe medir la calidad del profesorado, casi cualquier métrica produce el mismo resultado.

Los estudios sobre la calidad del profesor comienzan a partir de la línea de investigación sobre la función de producción educativa. En este campo, la literatura examina los resultados de los alumnos en términos de los recursos o factores educativos de producción disponibles para profesores, alumnos, escuelas y otras instituciones. Muchas de estas publicaciones se centraron en la calidad del profesor como determinante del rendimiento del estudiante. Desde el Informe Coleman, *Igualdad de oportunidades educativas* (1966), algunos investigadores han medido la calidad del profesorado en términos de exámenes sobre conocimientos generales o con la puntuación obtenida en los exámenes para obtener el certificado de aptitud pedagógica. El Informe Coleman concluye que hay una relación positiva entre el conocimiento del estudiante en materia de lectura y matemáticas y el resultado del profesorado en las pruebas de "habilidad verbal". FERGUSON (1991, 1998) demostró que los alumnos de aquellos profesores que obtuvieron mejores resultados en el Test para la Comprensión Auditiva del Lenguaje (TECAL), que mide las competencias lingüísticas y el conocimiento profesional del profesor, obtuvieron también mejores resultados en los exámenes de lectura.

ROCKOFF (2004) es precursor en el uso de datos de panel estadounidense para estimar efectos fijos del profesor, controlando por características de los estudiantes y efectos fijos del aula. Muestra que un incremento de un punto en la desviación estándar de la calidad del profesor aumenta la puntuación obtenida de los estudiantes en un examen nacional de matemáticas en 0.24 puntos. Con objetivos comparativos, podemos señalar que este efecto es similar al encontrado en el experimento aleatorizado STAR, como diferencia entre el efecto de estar en una clase pequeña (13-17 estudiantes) frente a una clase grande (22-25 estudiantes)¹². RIVKIN, HANUSHEK y KAIN (2005) estiman, analizando los

¹⁰ Los estudios estuvieron muy influidos por MCCALL (1923).

¹¹ La evidencia recopilada por el Education Endowment Fund apunta exactamente en la misma dirección.

¹² KRUEGER (1999).

datos de panel del *Texas School Project* de la Universidad de Texas en Dallas, que un incremento de una desviación estándar en la calidad del profesor tiene como límite inferior un aumento de 0.11 en la puntuación del alumno en pruebas cognitivas. METZLER y WOESSMAN (2010) usan un único conjunto de datos de resultados de lectura y matemáticas, de profesores y estudiantes de 6º grado de Perú, que les permite controlar por efectos fijos de estudiantes, profesores y asignatura. Los autores encuentran que un incremento de una desviación estándar en el conocimiento del profesor sobre la materia aumenta el resultado del alumno en un 10%. Por tanto, se puede concluir que hay un amplio consenso entre los investigadores sobre el impacto del conocimiento del profesor sobre los resultados de sus estudiantes. En general, los estudios que relacionan los conocimientos de los profesores con los conocimientos de los alumnos indican que un aumento de una desviación estándar del conocimiento del profesor, uno de los indicadores posibles de la calidad de la docencia, se traslada en 0,15 desviaciones estándar en los resultados de los alumnos en pruebas cognitivas.

En el estudio CHETTY y otros (2012) se estima que los estudiantes asignados a clases que tenían un profesor mejor, medido en este caso como un profesional de la educación con más de 10 años de experiencia, ganaban un 6,9% más en su empleo a la edad de 27 años que los asignados a profesores de menor "calidad". Una mejora de una desviación estándar en la calidad del profesor en un año generaría una ganancia de renta entre 170.000 y 214.000 dólares para una clase de 20 estudiantes. Con estos datos CHETTY y otros (2012) estiman que aumentar la calidad de un profesor desde el percentil 25 hasta el percentil 75 genera un valor presente neto de 320.000 dólares (calculado con una tasa de descuento del 3%) para una clase de 20 estudiantes. Esto significa que el salario de un profesor de calidad debería ser de 320.000 dólares al año, para que tuviera una retribución acorde a su contribución en términos de la mejora salarial futura de sus estudiantes. Este número da una señal clara sobre la enorme importancia que tiene la "calidad" del profesor en el aprendizaje de los alumnos, al menos en las primeras etapas escolares.

2.3. La importancia de las intervenciones tempranas

La evidencia científica sobre la importancia de intervenir en las fases tempranas del desarrollo del niño se acumula con rapidez¹³. Las diferencias en capacidades, tanto cognitivas como no cognitivas, entre individuos de distintos estratos sociales se generan muy pronto¹⁴. A los 5 o 6 años existen ya diferencias importantes en capacidades cognitivas entre niños de diferentes estratos socioeconómicos. Estas diferencias se mantendrán fundamentalmente estables durante los siguientes años. En general, cuanto más tarde comienzan las actuaciones sobre niños con dificultades, menor efecto tienen.

Es bien conocido que el nivel de las capacidades de los niños está muy correlacionado con el nivel educativo de los padres¹⁵. Pero el motivo puede ser la falta de capacidad de padres con pocos estudios para guiar el desarrollo de las habilidades cognitivas y no cognitivas de sus hijos, y no tanto la capacidad económica de la

familia. Una vez el niño accede al sistema educativo formal, las deficiencias en la formación de habilidades a una temprana edad le harán mostrar un rendimiento académico inferior a los niños de grupos socioeconómicos superiores. Por tanto, si el objetivo es mantener la equidad, las intervenciones públicas deben centrarse en la fase más temprana de la niñez. La equidad no se puede conseguir en la universidad. Aumentar la renta familiar a partir de subvenciones o reducciones de las tasas universitarias, cuando el joven ya está en la fase de ciclo vital de asistir a la universidad, prácticamente no tiene ningún efecto en la compensación de los bajos niveles de inversión previos. Las desigualdades hay que tratarlas en el origen y no en la universidad. Además, la elevada rentabilidad social de las actuaciones en edades tempranas justifica una intensa participación de la financiación pública.

Por tanto, cada vez es más evidente para los investigadores que lo que sucede antes de la edad preescolar es crítico¹⁶. El tipo de cuidados que reciben los niños durante los primeros tres años de vida es muy importante para ciertos efectos biológicos relacionados con la capacidad de atender y aprender. El concepto de *school readiness* no implica enfatizar el contenido académico antes de preescolar. *School readiness* se refiere a llegar a preescolar con un cerebro preparado y capaz de aprender. El aprendizaje comienza mucho antes de llegar a preescolar, pues la sinapsis comienza a producirse desde el nacimiento. El cerebro de un niño de dos años tiene casi el doble de conexiones neuronales que el de un adulto. Las conexiones que se refuerzan por la repetición no se pierden en el proceso de *neural pruning*¹⁷.

Pero además de los fundamentos biológicos de las intervenciones tempranas, existen experimentos que muestran la importancia de este tipo de actuaciones. Dos de los más conocidos son el Programa Preescolar de la Escuela Perry y el Programa Abecedario, que muestran cómo se pueden conseguir efectos a largo plazo de mejoras en habilidades cognitivas y no cognitivas, rendimiento académico y productividad laboral a partir de intervenciones tempranas. Por ejemplo, el programa Perry fue administrado a 58 jóvenes afroamericanos de Michigan entre 1962 y 1967. El tratamiento fue 2,5 horas de clase todos los días y 1,5 horas de visita a la familia cada semana. El Programa Abecedario estaba dirigido a jóvenes de familias desaventajadas nacidos entre 1972 y 1977. La media de entrada eran los 4,4 meses. La intervención era diaria. El programa Perry consiguió mejoras temporales del CI (desaparecieron a los cuatro años), pero el grupo tratado a los 14 años tenía mejores resultados académicos. Algunos autores argumentan que la explicación sería el efecto del programa sobre habilidades no cognitivas. Los individuos del grupo tratado del Perry (a los 40 años) y el Abecedario (a los 21 años) tienen mejores notas en pruebas académicas, mayores niveles educativos, requirieron una menor atención a través de educación especial, tenían mayores salarios, mayor probabilidad de tener una vivienda y menor probabilidad de estar en prisión que los individuos del grupo de control.

HECKMAN y otros (2009) muestran que la tasa de rentabilidad social anual del programa Perry se encuentra entre el 7% y el 10%. En términos de análisis coste-beneficio (suponiendo una

¹⁶ Para una aproximación desde el campo de la neurociencia ver SEBASTIÁN (2012)

¹⁷ Proceso neurológico que favorece un cambio en la estructura neuronal mediante la reducción de las conexiones sinápticas más débiles (en términos de su utilización) y permite mantener aquellas que generan una configuración sináptica más eficiente. Ver KNUDSEN y otros (2006) para una descripción de la convergencia en un conjunto de principios en diversas disciplinas como la economía, la psicología del desarrollo y la neurobiología sobre los potentes efectos del ambiente en los primeros años de la vida sobre el desarrollo de habilidades y la adquisición de conocimiento.

¹³ CURRIE (2001) ofrece una panorámica general.

¹⁴ Para una visión reciente de este tema véase CUNHA y HECKMAN (2010).

¹⁵ Entre los más recientes existen multitud de trabajos sobre la influencia del estatus socioeconómico en los resultados del estudio PISA. En el caso español, el trabajo de ANGHIEL y CABRALES (2010) proporciona la evidencia más convincente.

tasa de descuento del 3% y teniendo en cuenta el efecto de los impuestos necesarios para financiar el programa), el resultado es que de cada dólar gastado revierten a la sociedad entre 7 y 12 dólares en términos de valor presente.

En países menos desarrollados, la salud y la escolarización van de la mano. Para conseguir el objetivo de escolarización temprana, existe un requisito mínimo nutricional. Varios estudios han comprobado que los niños bien nutridos durante los primeros años de vida permanecen en el sistema educativo por más años y tienen un mayor aprovechamiento. El experimento INCAP de Guatemala (DAMON y GLEWWE, 2007) comprobó que suministrando atole (un suplemento nutricional) se producía un incremento medio de 0,6 años de escolarización en las aldeas tratadas frente a las aldeas de control. En un experimento muy interesante KREMER y MIGUEL (2004) muestran cómo la eliminación de los gusanos intestinales en niños de Kenia, además de mejorar su salud, tiene como efecto colateral un aumento de la asistencia a clase. Además, el procedimiento es muy efectivo en términos del coste que supone conseguir un año más de escolarización.

Por tanto, el objetivo de conseguir incrementar sustancialmente la formación en edades tempranas debe acompañar las medidas de oferta (centros, profesores, etc.) con medidas de demanda como la señalada. La ratio de beneficio sobre coste es de 4,6, a una tasa de descuento del 3%. El programa PIDI (Proyecto integral de desarrollo infantil) es otro ejemplo de intervenciones efectivas. Destinado a niños entre 6 y 72 meses, es un caso de combinación de medidas nutricionales y educativas en educación temprana. DAMON y GLEWWE (2007) calculan un ratio beneficio coste de 3,3 (3% de tasa de descuento) o 2,2 (al 6% de descuento).

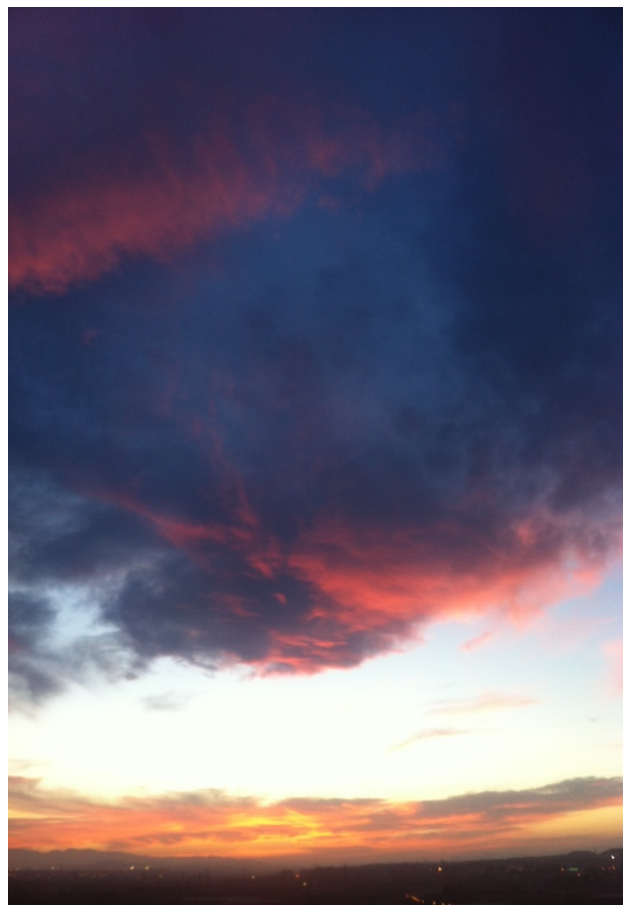
2.4. Otras evaluaciones

Hay muchos otros temas muy interesantes en el campo de la evaluación educativa. Por ejemplo, un tema que recibe mucha atención es la importancia de los compañeros de clase en el aprendizaje. El análisis de la varianza en CHETTY y otros (2012) muestra que la identidad de la clase tiene un efecto significativo sobre los salarios cuando los jóvenes se incorporan al mercado laboral. Los estudiantes que fueron aleatoriamente asignados a clases donde sus compañeros en el último año de guardería tenían mejores resultados; ganaban mayores salarios y tenían una probabilidad de asistir a la universidad superior a los que acabaron asignados a clases de menor calidad. El efecto de tener compañeros con más conocimientos sobre las pruebas cognitivas se va reduciendo en los años posteriores, pero reaparece en el mercado laboral. Los autores argumentan que la explicación es la persistencia de las ganancias en aspectos no cognitivos.

3. La situación española: pánico a la evaluación, falta de transparencia y dominio ideológico

Ciertamente la situación de la evaluación del sistema educativo en España es mucho peor que la observada en otros países que aceleran hacia política orientadas por la evidencia, como Estados Unidos o el Reino Unido. Recuerdo que a finales de los 90 trabajaba frecuentemente en la cuestión de la inserción laboral de los universitarios. Cuando contactábamos con los rectorados de diferentes universidades para solicitarles la colaboración en el envío de cuestionarios a sus graduados para analizar su situación laboral, la respuesta solía ser no; o sí, pero con condiciones: nunca se realizaría una comparación de los indicadores de inserción laboral entre universidades. El enconado enfrentamiento sobre la evaluación de la asignatura de Religión o la eliminación de Edu-

cación para la Ciudadanía durante la discusión de la última reforma educativa, o la ausencia de debate basado en evidencia sobre aspectos sustanciales del sistema anterior son muestras del deprimente estado en que se encuentra la relación entre la política educativa española y la evidencia científica en el campo de la educación. En muchos casos, cuando se propone hacer evaluaciones que permitirían obtener evidencia empírica se suscita una gran polémica que frecuentemente ha frustrado el avance del conocimiento. Peor aún; en ocasiones se han financiado programas para los cuales existía evidencia de su escasa o nula efectividad, sin tener en cuenta aspectos importantes sobre su implementación. Este es el caso del programa Escuela 2.0 que tenía como objetivo financiar un portátil para cada alumno de 5º y 6º de Primaria y 1º de ESO de 14 comunidades autónomas (Madrid, Valencia y Murcia no participaron). El coste del programa ascendió a 600 millones de euros y se inició en 2009, cuando la crisis económica ya había aumentado significativamente el coste de oportunidad de los fondos públicos. Hay docenas de estudios sobre el efecto de CAL (“computer assisted education”) en el aprendizaje. Los resultados dependen de las condiciones del programa y la metodología estadística empleada. Uno de los más recientes, y metodológicamente mejor realizados, es BARRERA-OSORIO y LINDEN (2009). En este trabajo los autores presentan la evaluación de un programa para enseñar idiomas en escuelas públicas de Colombia utilizando ordenadores que, en este caso, fueron financiados por el sector privado. Seguramente este procedimiento de financiación favoreció el interés de los financiadores por evaluar apropiadamente sus efectos. Se trata de un experimento aleatorizado que duró dos años y que incluyó una muestra de 97 colegios y 5.201 estudiantes. Los resultados de la investigación muestran que el programa tuvo muy poco efecto sobre los resultados de los estudiantes, en términos de conocimientos adquiridos. Los autores señalan que este tipo de programa “implementado de forma masiva a nivel nacional, no tiene efecto en los resultados de los estudiantes”.



Como en cualquier experimento, es lógico plantearse la validez externa en otros contextos educativos de los resultados obtenidos en el caso específico de Colombia. Pero ante la ausencia de evaluaciones en un país concreto, es mejor usar esta evidencia como base para una reflexión sobre la validez externa de los resultados aplicados a otros países, que tomar decisiones sin ninguna base científica. Además el *Education Endowment Fund* resume de forma muy precisa el conocimiento acumulado por las evaluaciones disponibles sobre el impacto de las tecnologías digitales en la clase. La evidencia sugiere que el aprendizaje individualizado con componentes tecnológicos (como los programas basados en la provisión individualizada de ordenadores personales a cada estudiante, como el Escuela 2.0) no tiene un efecto significativo, mientras el aprendizaje en pequeños grupos, o el uso colaborativo de la tecnología, parece tener un impacto positivo. Además, “hay evidencia clara de que es más beneficiosa en áreas como la escritura, que para aprender a deletrear, o en matemáticas, más que para la resolución de problemas”.

No obstante en España algo se empieza a mover. El Instituto Nacional de Evaluación Educativa (INEE) parece haber tomado recientemente la línea que adoptó su homólogo norteamericano a favor de generar conocimiento para poder realizar intervenciones educativas basadas en la evidencia. El INEE recopila y pone a disposición de los investigadores los microdatos de los estudios internacionales de evaluación, lo que garantiza una mejora sustancial en el conocimiento de los factores determinantes del rendimiento educativo en España.

4. Conclusiones

La educación es una de los factores más importantes para la consecución de los objetivos básicos de cualquier sociedad. Desde esta perspectiva, parece lógico que las intervenciones y las reformas educativas sean evaluadas de forma rigurosa y científica, evitándose los aprioris basados en planteamientos ideológicos y las posiciones que no tengan un soporte en la evidencia empírica. Por suerte, en los últimos años se ha acumulado con rapidez un conjunto amplio de evidencia científica sobre lo que funciona, y lo que no funciona, en el campo de las intervenciones educativas. Este conocimiento se utiliza rutinariamente para orientar políticas en otros países. Por desgracia en España todavía estamos muy lejos de adoptar esta aproximación basada en la evidencia, aunque últimamente se están dando pasos en esta dirección.

Referencias bibliográficas

ANGHEL, B. y CABRALES, A. (2010). *Los determinantes del éxito en la educación primaria en España*. Madrid: FEDEA

ANGRIST, J. (2004). “American education research change tack”. *Oxford Reviews of Economic Policy*, vol. 20, nº 2, pp. 198-212.

BARRERA-OSORIO, F. y LINDEN, L. (2009). “The use and misuse of computers in education: evidence from a randomized experiment in Colombia,” *Policy Research Working Paper* 4836, The World Bank.

BOUGUEN, A. y GURGAND, M. (2012). “Randomized controlled experiments in education”. *European Expert Network on Economics of Education Analytical*. Report 11, European Commission.

CHETTY, R., FRIEDMAN, J., HILGER, N., SAEZ, E. SCHANZENBACH, D. y YAGAN D. (2011). “How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star”. *Quarterly Journal of Economics*, vol. 126, nº 4, pp 1593-1660.

COOK, T. (2001). “Sciencefobia: Why education researchers reject randomized experiments”. *Education Next*, Fall, pp. 63-68.

CUNHA, F. y HECKMAN, J. (2010). “Investing in our young people”. *NBER Working Paper* 16201.

CURRIE, J. (2001). “Early childhood education programs”, *Journal of Economic Perspectives*, vol. 15, nº 2, pp. 213-238.

DAMON, A. y GLEWWE, P. (2007). “Three Proposals to Improve Education in Latin America and the Caribbean: Estimates of the Cost and Benefits of Each Strategy: A Summary”. *Estrategia y Negocios*, Agosto.

HANUSHEK, E. y WOESSMANN L. (2008), “The role of cognitive skills in economic development”. *Journal of Economic Literature*, vol. 46, nº 3, pp. 607-668.

HANUSHEK, E. y WOESSMANN L. (2010). *The high cost of low educational performance: the long run impact of improving PISA outcomes*. Paris: OECD PISA Program.

HECKMAN, J., MOON, S. H., PINTO, R., SAVALYEV, P. y YAVITZ, A. (2009). “The rate of return to the HigScope Perry Preschool program”, *Journal of Public Economics*, nº 94, pp. 114-128.

KREMER, M. y MIGUEL, E. (2004). “Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities”. *Econometrica*, vol. 72, nº 1, pp. 159-217.

KRUEGER A. (1999). “Experimental Estimates of Education Production Functions”. *Quarterly Journal of Economics*, vol. 115, nº 2, pp 497-532.

KNUDSEN, E.I., HECKMAN, J., CAMERON, J. y SHONKOFF, J.P. (2006), “Economic and neurobiological and behavioral perspectives on building America’s workforce”. *Proceedings of the National Academy of Sciences*, vol. 103, nº 27, pp. 10155-10162.

MCCALL, W. (1923). *How to experiment in education*. New York: The Macmillan Company.

METZLER, J. y WOESSMANN, L. (2010). “The Impact of Teacher Subject Knowledge on Student Achievement: Evidence from Within-Teacher Within-Student Variation”. *IZA Discussion Paper*, nº. 4999.

RIVKIN, S., HANUSHEK, E. y KAIN, J. (2005). “Teachers, schools and academic achievement”. *Econometrica*. Vol 73, nº 2, pp 417-458.

ROCKOFF, J. (2004). “The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data”. *American Economic Review*, nº. 94 (2), pp. 247-252.

ROCKOFF, J. (2009). “Field experiments in class size from the early twentieth century”. *Journal of Economic Perspectives*, vol. 23, nº 4, pp 211-230.

SEBASTIÁN, N. (2012). “Neurociencia cognitiva del desarrollo: el periodo preescolar”. *Participación Educativa*. Vol. 1, nº 1-II, pp. 34-38.

El autor

José García Montalvo

Es Catedrático de Economía de la Universitat Pompeu Fabra. Licenciado en Ciencias Económicas por la Universitat de València (1987). Premio extraordinario de licenciatura (Universitat de València) y Primer Premio Nacional Fin de Carrera (1988, Ministerio de Educación y Ciencia). Doctor en Economía por la Universidad de Harvard (1993). En 2008 fue galardonado por su labor investigadora con la distinción ICREA Academia en su primera edición. En 2010 el Consejo Social de la UPF le entregó el premio a la Transferencia de Conocimiento. Ha sido consultor de la Unión Europea, la OCDE, del Banco Mundial y el Banco Interamericano de Desarrollo. Ha sido profesor invitado por el departamento de Economía de la Universidad de Harvard (1996) y trabajó durante un año como investigador del Programa de Tecnología y Política Económica de la Kennedy School of Government (1995). Ha desempeñado labores de gestión universitaria como Director del Departamento de Economía y Empresa (2007-2010) y Vicerrector de Política Científica (2011-2013) de la Universitat Pompeu Fabra. Ha publicado doce libros y más de cien artículos científicos.

